A Shortest Path in an Experimentally Built Semantic Network

Maciej Godny[1], Izabela Gatkowska[1], Wiesław Lubaszewski[1]

[1] Jagiellonian University in Kraków, Poland
{m.godny, i.gatkowska, w.lubaszewski}@epi.uj.edu.pl
http://www.klk.uj.edu.pl/klk/en

Abstract. From linguistic perspective a semantic network built by the free word association experiment is a structure to explain a lexical meaning Deese (1965) and a human associating mechanism Clark (1970). It is interesting to find if network's path analysis may contribute to both hypotheses. Due to the nature of an experiment we restrict the analysis to shortest paths which can replace the stimulus – response connection, that is the only pairs produced by an experiment. We shall analyze a formal properties and a semantic consistency of the shortest path. We shall treat the network as an undirected weighted graph. All results are based on Polish experimental semantic network (Gatkowska, 2017).

1. Experimental Semantic Networks

One can observe the use of an experimentally built semantic networks in cognitive oriented NLP. For example network can be treated as a norm to evaluate a performance of algorithms which generate associations from text (Rapp, 2002; Wandmacher et al., 2008; Korzycki et al., 2017). One may also look into the network to find an information which is not lexically present in a text, e.g. *barks/barking* in lexical level implies a *dog* in semantic level (Lubaszewski et al., 2017). Finally, one may look at the network as a source to study human association mechanism.

Experimental semantic networks, e.g. English – Kiss (Kiss et al., 1973, Nelson et al., 1999), Dutch – DeDeyne and Storms (2008), Polish – Gatkowska (2015, 2017) are built by a use of the free word association test Kent-Rosanoff (1910), in which the tested person responds with a first word associated with a stimulus word provided by a researcher. The resulting network is a directed weighted graph. Each connection between two nodes has a direction, always from stimulus to response. Each connection has an association strength which is a number of participants, who are associated particular response with a particular stimulus, e.g. home – house 24 in EAT, i.e. first ever English experimentally built semantic network (Kiss et al., 1973).

Experimental semantic networks differ from manually built taxonomic networks as e.g. WordNet or CYC. Steyvers and Tenembaum (2005) had compared an experimental network Nelson et al. (1999) to WordNet and Roget's Thesaurus. The most important fragment of comparison which refers to an average number of connections for network node $<k>$, and average shortest path length L, is shown in Table 1.

Table 1. Comparison of experimental and taxonomic networks

	Nelson Undirected	Nelson Directed	Roget	WordNet
$<k>$	22.0	12.7	1.7	1.6
L	3.04	4.27	5.60	10.56

One may explain the difference in connection numbers $<k>$ in following way. A taxonomic network uses mainly paradigmatic relations (subordinate – superordinate; part – whole) to link the nodes. Participant of the free word association experiment is not restricted and may use any type of semantic dependency to link the two words, which was observed by Deese (1965), Clark (1970). This answering freedom produces more connections per node in an experimental network, which means that description of a node meaning is richer, because we have multiple syntagmatic dependencies Clark (1970), which are properties like color, size, purpose etc. Gatkowska (2017). If we look at the *L* parameter we may say that the path length in taxonomic networks reflects a taxonomy depth, which does not exist in an experimental network.

2. Statistic Structure of an Experimental Semantic Network

The analysis of the formal properties of a experimentally built semantic network which was based on American English network (Nelson et al. 1999) was made by Steyvers and Tenembaum (2005). The comparative analysis of Dutch and English networks was made by DeDeyne and Storms (2008). Both analyses were performed on reduced networks, i.e. authors truncated all connections with association strength equal to 1, which means connections produced by a single person in the experiment. Both analyses considered a network as an unweighted graph. We shall use the criteria introduced by Steyvers and Tenembaum (2005) to analyze the Polish network (Gatkowska, 2017) and to compare Polish network to English and Dutch networks.

The criteria are as follows: n = the number of stimuli nodes; L = the average shortest path length, computed for each pair of all network nodes; D = the diameter of the network, i.e. the length of the longest path in all network shortest paths; <k> = the average number of incoming connections in directed network and in undirected network the <k> is the average number of connections to all neighbouring nodes. The Table 2. shows that compared networks are of different size. Nelson used 5018 stimuli, De Deyne 1424, Gatkowska 316 but the network structures seems to be similar.

Table 2. Network structure comparison.

	Nelson et al., 1999		DeDeyne, 2008		Gatkowska, 2017	
	directed	undirected	directed	undirected	directed	undirected
n	5018	5018	1424	1424	316	316
L	4.27	3.04	3.43	2.46	3.81	3.99
D	10	5	9	4	7	6
<k>	12.7	22.0	17	29	9.52	49.81

The differences in size of <*k*> are caused by differences in number of persons who responded to a particular stimulus i.e. 100–120 in Nelson experiment, 82–197 in DeDeyne and Storms, and 861–893 in Gatkowska.

3. The aim of the paper

Both papers cited above admitted that it would be interesting to analyze a shortest path in an experimental semantic network, but both did not analyze the shortest path in details. The aim of this paper is to present a formal properties and semantic consistency of the shortest path in an experimentally built semantic network. We shall make four assumptions.

First, to perform a shortest path analysis we have to refer to linguistic analyses of a network produced by the free word association experiment. Those analyses made from a linguistic point of view were based on assumption that each connection between stimulus and response is based on a semantic dependency between meaning of the stimulus word and the response word: Deese (1965), Clark (1970). This implies that a shortest path analysis may show unknown semantic properties of a network.

Second, we shall look only for shortest paths which may replace a stimulus – response direct connection. This kind of path may explain those stimulus – response connections that are not explicable by semantic relations, e.g. *baranina* – *wełna* (mutton – wool). This kind of path may enrich the meaning definition of the word that is a node of a network and may bring a data to study of human associating mechanism.

Third. To perform an analysis we shall treat a semantic network as an undirected graph. We know from linguistic analyses, e.g. Clark (1970), Gatkowska (2017) that direction of semantic relation between two words is independent from stimulus – response direction, i.e. if we have two pairs of different directions e.g. *bulldog*

→ *dog* and *dog* → *bulldog*, we can find that there is just one semantic relationship which goes from subordinate *bulldog* to superordinate *dog* (hyponymy).

Finally, we shall acknowledge that a semantic network built by the free word association experiment is a weighted graph, because each stimulus – response connection has an experimentally assigned association strength.

4. Experimental Semantic Network of Polish

Regarding network similarities shown in Table 2. Our analysis of a shortest path will base solely on Gatkowska (2017). In the experiment Gatkowska used 322 stimuli but we excluded from our analysis 6 expressions which were used as stimuli, i.e. *na świat, nie ma, za wsią, wuja Toma, do papieru, z dziurami*. We exluded also empty answers, created by participants who refused provide an answer to a particular stimuli. Finally, we excluded all answers with spelling error. Due to analyses made by Deese (1965), Clark (1970), Gatkowska (2015, 2017), which show that semantic quality of the stimulus – answer connection does not depend on association strength, we included all answers provided by a single person. The resulting network is described in Table 3.

Table 3. Polish network structure.

stimuli	nodes	connect dir	connect undir
316	12 182	41815	36 460

Due to nature of the free word association experiment in the undirected network the number of connections is equal to the number of outgoing connections which are present in directed network.

5. Shortest Path in an Experimental Semantic Network – Formal Properties

To perform an analysis we shall treat a semantic network as the undirected weighted graph. The shortest path between pair of nodes in the network means a path which starts with one node of the pair (start node) and ends with a second node of the pair (end node) and its path weight (the sum of weights of the path connections) is the smallest of all possible paths for a particular pair of nodes. For example in experimental semantic network for nodes *baranina – wełna* (mutton – wool) the path *baranina – baran – owca – puszysty – wełna* (mutton – ram – sheep – fluffy – wool) with path weight of 35 is shoter than path *baranina – mięso – lama – wełna* (mutton – meat – lama – wool) with path weight 39. If the two paths have the same path weight both are considered as a shortest path. In other words, it is possible in the semantic network that we can have more than one shortest path for a particular pair of nodes.

5.1 The Weight System

The association strength S is a number of test respondents who linked a particular response with the stimulus, for example if *dom – dach* (house – roof) were linked by 17 persons then S=17. One may convert an association strength into weight W using a total number of responses given to a particular stimulus T and it's relation to S. We shall use two methods of weight computing. The traditional one abbreviated as W1 (Kiss et al., 1973) where weight is computed as S/T; in our example of *dom – dach* where T = 872 and S = 17 the weight W1 = 0.019 (17/872). The second method abbreviated as W2 is computed as T/S., which means that for connection *dom – dach* W2 = 51.29 (872/17). In the case when two words are reversely a stimulus and a response, e.g. *kapusta* → *kiszona* (cabbage – saured) and *kiszona* → *kapusta* we have bidirectional connection with separate weights for each connection. In the undirected network we have to use only one weight. We decided to use a stronger one.

5.2 Path number

We shall use both weights (i.e. W1, W2) to find if the number of shortest paths which are in the network depends on weight. To test both weights we use three different values of association strength S between two nodes: $S \geq 1$, $S \geq 2$ and $S \geq 10$. Figure 1. shows that path number in fact depends on S, because for both weights the number of paths differs substantially only with $S \geq 1$.

If we consider fact, that in the experimental semantic network the weight of connection between the two nodes is set empirically, then we may add a new, more restrictive criterion in our analysis. Which means that we may look for the shortest paths which have the path weight smaller than the weight of direct connection between stimulus and response. Results of such restricted shortest path extraction are shown in Figure 2.

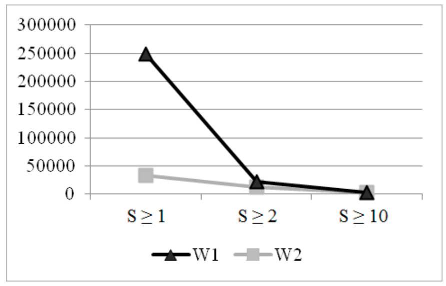

Figure 1. Path number.

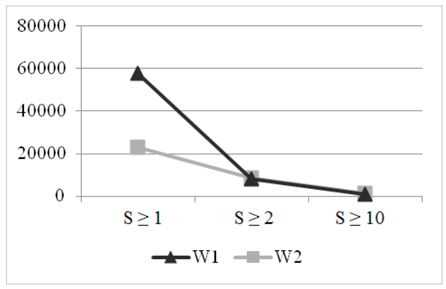

Figure 2. Path number – restricted.

As we can see the introduced restriction that relates path weight to the weight of direct connection makes the number of paths substantially smaller. We can also observe that both weights, i.e. W1 and W2 produce a similar paths number with S >= 2 and S ≥ 10.

To summarize one may say that for both weights W1, W2 there is a minimum value of S which produces a similar number of paths.

5.3 Path length

We shall treat a path length as a number of nodes in the path, including end and start nodes. We analyzed the path length for weights W1 and W2 with association strength S of values ≥ 1, 2 and 10. The relation between weight values and path length shows the Figure 3.

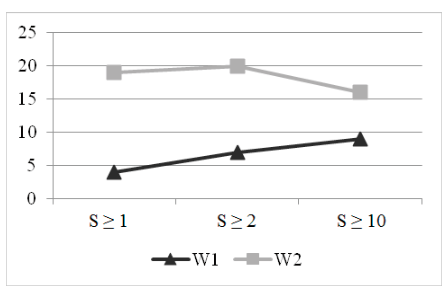

Figure 3. Maximum path length.

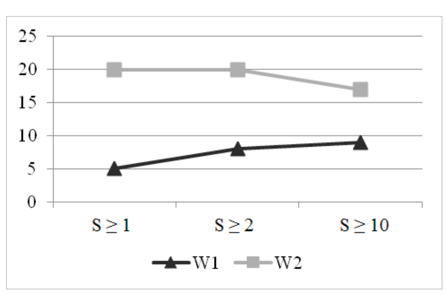

Figure 4. Maximum of the restricted path length.

Results show that the raise of weight value results in maximum of the path length. In the case of W2 we can observe that the maximum length of the path decreases. In the case of W1 we can see the opposite direction, the maximum length of the increases rapidly.

The detailed results are shown in Table 4, where pl means path length counted in nodes; weights W1 and W2 are counted as described above; S values are of ≥ 1, 2 and 10. Table shows a path number as a percentage of path total.

We can see in Table 4, that weight W1 produce much shorter paths than W2 – the longest path produced by W1 consists of 9 nodes compared to 20 for W2. For both W1 and W2 we can observe that the raise of the weight value results in increase of number of the short paths, i.e. of 3–4 nodes. But for W1 the number of paths longer that 4 also raises.

Table 4. Detailed unrestricted path length.

pl	S ≥	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
W1	1	16,7	83,3																
	2	22,9	48,1	22,4	5,6	0,3													
	10	31,5	35,1	15,3	11,3	4,3	2,0	0,5											
W2	1	15,8	17,4	10,8	13,2	11,2	10,1	8,2	5,2	3,5	2,0	1,2	0,7	0,4	0,2	0,1	0,1	0,1	
	2	23,1	21,9	10,6	11,8	8,9	8,3	6,1	3,8	2,5	1,3	0,8	0,5	0,2	0,2	0,1	0,1	0,1	0,1
	10	34,0	25,3	9,5	10,4	6,1	4,9	3,9	2,6	1,5	1,0	0,5	0,3	0,1	0,2				

Now we shall repeat path length analysis using restriction that says that path weight (the sum of weights of the path connections) should be smaller than the weight of the direct connection. Results are presented in Figure 4.

As we can see the condition does not change the observed earlier relation between the weight value and the path length. The detailed results are shown in Table 5. where path number is a percentage of path total.

Table 5. Detailed length of the restricted path.

pl	S ≥	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
W1	**1**	25,0	68,9	6,2															
	2	29,8	46,6	18,8	4,7	0,2	0,1												
	10	39,4	33,9	13,7	8,9	3,1	0,8	0,2											
W2	**1**	13,9	16,5	11,5	13,6	11,7	10,6	8,5	5,5	3,6	2,0	1,3	0,7	0,4	0,2	0,1	0,1	0,1	0,1
	2	21,6	21,7	11,7	12,0	9,2	8,7	6,3	3,7	2,4	1,2	0,8	0,5	0,2	0,1	0,1	0,1	0,1	0,1
	10	42,5	25,2	10,7	8,6	4,3	3,5	2,6	1,3	0,5	0,2	0,2	0,2	0,1	-	0,1			

Table 5. shows that the restricted extraction of the shortest path reduces path numbers (see Figure 2) but does not change relation between weight and path length.

5.4 Path Structure

5.4.1. Start Node and End Node structure

Due to lack of space we shall not provide an analysis of POS patterns which may occur in paths extracted from experimental semantic network. Where pattern means part of speech structure correlated with a path length. We shall restrict our analysis to the analysis of start and end nodes of the path regarding the three main parts of speech. We performed our analysis using path connection weights W1 and W2 with S ≥ 1. We show comparative results for both unrestricted and restricted shortest path extraction. Results are shown in Table 6. where N stands for Noun, V for Verb and A for Adjective.

Table 6. POS structure of the path.

	Unrestricted		restricted	
Sn- -En	W1	W2	W1	W2
N- -N	48%	48%	51%	51%
N- -V	2%	2%	3%	3%
N- -A	15%	15%	15%	15%
V- -N	4%	3%	4%	3%
A- -N	26%	26%	23%	23%
A- -A	3%	3%	3%	3%

First, looking at Table 6., we can find that the percentage of start node POS related to the end node POS does not depend on connection weight. One can suppose that the observed stability is an effect of stabile semantic dependency between the stimulus and response. Secondly, the observed results may support Deese's (1965) hypothesis that the set of responses provided by participants of the free word association experiment constitute the meaning of the stimulus word. And we know the word meaning is and should be stable.

5.4.2 Intermediate nodes in a path

In the case of intermediate nodes we are going to find all nodes which do not enter in direct connection with a start node in the network. In other words, nodes without a direct connection to a stimulus'. Such nodes are crossing the frontier of a subnet created by direct connections to a stimulus, which – according to Deese (1965) – define a stimulus meaning. We can mark automatically those intermediate nodes in a path. Results for both weights W1 and W2 for unrestricted path extraction are presented in table 7. Where we can see the number of paths which contain at least a single node which is not directly connected to start node are a large fraction of the total path number for both weights.

Table 7. Paths with intermediate nodes without a direct connection to a stimulus.

Weight	S ≥ 1	S ≥ 2	S ≥ 10
W1	70%	59%	45%
W2	66%	56%	57%

Having a global statistics of the paths extracted by an unrestricted algorithm we have to do an analysis of nodes which introduce those frontier crossing connections. We shall present a part of speech analysis – see table 8.

Table 8. A part of speech analysis.

	W1			W2		
	S ≥ 1	S ≥ 2	S ≥ 10	S ≥ 1	S ≥ 2	S ≥ 10
N	65%	73%	78%	83%	83%	83%
V	1%	2%	2%	1%	1%	1%
A	12%	20%	15%	15%	14%	15%
other	21%	5%	6%	1%	1%	1%

Table 8 shows that the frontier crossing connections are introduced mainly by nouns and other which are mainly proper names and pronouns. But such an observation does not explain how those frontier crossings relate to the meaning of a start node. In the next section we perform an analysis of the problem.

6. Semantic Consistency of a Path

As has been told above each pair of directly connected nodes is semantically connected due the nature of the free word association experiment which produces semantically related stimulus – response pairs. If we look at a path between two directly connected nodes where the start node is a stimulus and the end node is an answer we may find that some intermediate nodes in a path enter into semantic relationship with the start node and we may also find that some nodes are not related to the start node. If all intermediate nodes in a path are semantically related to the start node we may say that a path is semantically consistent, e.g. path *boleć – noga – ręka* (pain – leg – hand) where *leg* and *hand* enter the same 'location' relationship to *pain*. If there is an intermediate node in a path which does not enter in semantic relationship with the start node the path is not consistent, e.g. in the path *kapusta – zielona – roślina – kwiat* – roża* – czerwona,* (cabbage – green – plant – flower* – rose* – red) only two intermediate nodes enter in semantic relationship with the start node, i.e. *cabbage* 'color' *green* and *cabbage* 'hyponymy' *plant*. The inconsistent nodes are marked by asterisk (*). One may say that the inconsistent path in some point crosses the line which separates a subnet which defines a meaning of a start node as in Figure 5. where start and end nodes are darkened.

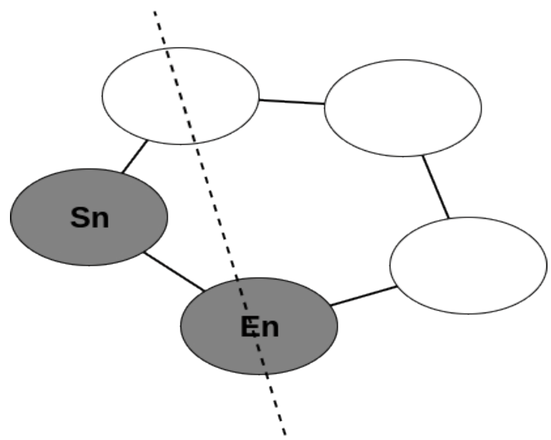

Figure 5. Subnet which defines a meaning of a start node.

One may also observe in our example of an inconsistent path that the intermediate nodes which do not enter into semantic relationship with the start node are not connected directly to the start node in the network. It would be interesting to find if those marked nodes caused that the shortest path extracting algorithm crossed the line which separates meanings in the network.

Unfortunately, analysis of path consistency should be done manually. We have no proper lexical tools to perform analysis automatically because WordNet uses only paradigmatic lexical relationships, i.e. hyponymy, meronymy and so on. On the other hand the FrameNet which describes syntagmatic dependencies does not use lexical relationships. Therefore we shall present results of manual analysis of shortest paths extracted by unrestricted algorithm for weight W2 with value of S ≥ 10, which means an analysis of 3260 paths. But regarding the similarities between W1 and W2 already shown one may think that this analysis is sufficient.

6.1 Consistency vs. directness of connection

Results of the path semantic consistency analysis are shown in table 9 where indirect stands for paths that contain at least one node which do not enter into direct connection with start node, and direct stands for paths with all nodes directly connected to a start node.

Table 9. Consistent vs. inconsistent paths.

	consistent	**inconsistent**
total	1765	1495
indirect	317	1161
direct	1448	334

First we can see in the table that semantically consistent paths share only 54% of extracted paths. Table 9 also shows that there are 18% of consistent paths which contain at least one node which do not enter into direct connection with a start node and there are 22% of inconsistent paths which are built with nodes directly connected to a start node. This implies that both direct and indirect connectivity do not fully explain path consistency. Finally, the fact that 82% of consistent paths were built from nodes directly connected to a start node rises doubts if those paths may expand Deese's definition of an associative meaning.

6.2 Path Consistency vs. Path Length

Table 10 presents results for 1765 semantically consistent paths evaluated manually, where column *Pl* (path length) presents node number in a path; columns next to the right show a number of indirect nodes in a path *In*, i.e. nodes which are not directly connected to the start node by a free word association experiment.

Table 10. Number of indirect nodes in semantically consistent paths.

Pl \ In	0	1	2	3	4	Total
3	973	0	0	0	0	973
4	381	180	0	0	0	561
5	61	46	16	0	0	123
6	32	36	21	1	0	90
7	1	3	4	4	0	12
8	0	2	1	1	**1**	5
9	0	1	0	0	0	1
	1448	**268**	**42**	**6**	**1**	**1765**

As we can see in the table 10 the path length is determined by a first node, which is indirectly connected to a start node (stimulus). Secondly, we may observe that the consistent paths which has a length greater than 5 share 6% of consistent paths total which means that there is no need for arbitrary fixed path length frequently used in analyses e.g. Steyvers and Tanembaum (2005).

Table 11. Number of indirect nodes in semantically inconsistent paths.

Pl \ In	0	1	2	3	4	...	Total
3	136	0	0	0	0	...	136
4	128	134	0	0	0	...	262
5	49	76	61	0	0	...	186
6	17	65	93	74	0	...	249
7	4	29	43	66	44	...	186
8	0	7	12	39	57	...	155
9	0	3	2	11	32	...	126
...	...	...	...	...	...	...	...
	334	**314**	**214**	**193**	**147**	**...**	1495

The table 11 shows that inconsistent paths are longer than consistent ones presented in table 10. We can also see that the path length is correlated with number of nodes indirectly connected to a start node. Finally, we can not say that path inconsistency is introduced by nodes indirectly connected to a start node, there are 334 (22%) without indirectly connected nodes. We think that a real reason of path inconsistency is a semantic value of direct connection between start node and end node, that is value of a connection between stimulus and response. The next section will bring a sketch of the problem.

6.3 Inconsistency Analysis

At this stage of the research we are not able to explain fully what makes a path a consistent or inconsistent. But we have found the following main situations in which the path may appear semantically inconsistent.

- a path is always inconsistent, when start node and end node are constituents of an idiom or multipart word, e.g. *czarny – owca* (black – sheep) which constitute the idiom *black sheep*. For this example we obtained paths like: *czarny – lęk – owca* (black– fear – sheep) or *czarny – kot – zwierzę – baran – owca* (black – cat – animal – ram – sheep):

- a path is frequently inconsistent when start node and end node enters in relation hyponymy (subordinate → superordinate), e.g. *owca – zwierzę* (sheep – animal) which produces paths like: *owca – owoc – zwierzę* (sheep – fruit – animal) or *owca – lew – zwierzę* (sheep – lion – animal).
- a path is frequently inconsistent, when start node and end node enter in relation meronymy (part – whole), e.g. *dach – dom* (roof – house), which produces paths like: *dom – szyba – dach* (house – glass – roof) or *dom – dym – dach* (house – smoke – roof).
- a path is frequently inconsistent, when the start node or end node is a feature that belongs to many objects, e.g. *duży – dom* (big – house) *duży – młot – dom* (big– hammer – house) or *duży – kozioł – dom* (big – goat – house).

Finally, we have to stress that our analysis is made to signal rather than solve the problem of the path semantic consistency.

7. Conclusion

The analyses presented in the paper are preliminary ones but they show that a pure statistical approach may not explain all properties of an experimental semantic network fully. But even such preliminary analysis suggests that only small fraction (18%) of consistent paths may bring an extension to a word meaning defined in the network. This implies also that a consistent path analysis may bring valuable data to study a semantic driven human association mechanism.

It seems to be clear that there is a need for further investigation of the semantic consistency of a path in the network to find if there is possible to create an algorithm which can extract only consistent paths in a network.

References

Clark, H.H. (1970). Word associations and linguistic theory. J. Lyons (ed.), *New Horizons in Linguistics*, Harmondsworth: Penguin Books, Harmondsworth.

DeDeyne, S., Storms, G. (2008). Word associations: network and semantic properties, *Behavior Research Methods*, *40*(1), 213-231.

Deese, J. (1965). *The structure of Associations in Language and Thought*, Baltimore, John Hopkins University Press.

Gatkowska, I. (2017). *Eksperymentalna sieć leksykalna języka polskiego* (*An Experimental Lexical Network of the Polish Languag*e), Kraków, Wydawnictwo Uniwersytetu Jagiellońskiego.

Gatkowska, I. (2015). Empiryczna sieć powiązań leksykalnych, *Polonica*, Volume 35, 155-178.

Kent, G. & Rosanoff, A.J. (1910). *A study of association in insanity. American Journal of Insanity*, *67*(2), 317-390.

Kiss, G.R., Armstrong, C., Milroy, R. et al. (1973). An Associative thesaurus of English and its computer analysis. A.J. Aitken, R.W. Bailey, N. Hamilton-Smith (eds.), *The computer and Literary Studies*, Edinburgh: University Press.

Korzycki, M., Gatkowska, I., Lubaszewski, W. (2017). Can the Human Association Norm Evaluate Machine-Made Association Lists? B. Sharp, F. Sèdes & W. Lubaszewski (eds.), *Cognitive Approach to Natural Language Processing*, 21-38.

Lubaszewski, W., Gatkowska, I. & Haręza M. (2015). Human association network and text collection. B. Sharp & R. Delmonte (eds.), *Proceedings of the 12th International Workshop on Natural Language Processing and Cognitive Science 2015*, Berlin, De Gruyter, 101-114.

Lubaszewski, W., Gatkowska, I., Godny, M., (2017). How a Word of a Text Selects the Related Words in a Human Association Network. B. Sharp, F. Sèdes & W. Lubaszewski (eds.), *Cognitive Approach to Natural Language Processing*, 41-62.

Nelson, D.L., McEvoy, C.L., & Schreiber, T.A. (1999). *The University of South Florida word association norms.* Retrieved from http://w3.usf.edu/FreeAssociation. Access date: September 25, 2017.

Rapp, R. (2002). The computation of words associations: comparing syntagmatic and paradigmatic approaches. *Proceedings of the 19th International Conference on Computational Linguistics*, vol. 1, 1-7, Taipei, Taiwan, August 24-September 1.

Wandmacher, T., Ovchinnikova, E. & Alexandrov, T. (2008). Does latent semantic analysis reflect human associations? *Proceedings of the ESSLLI Workshop on Distributional Lexical Semantics*, Hamburg, Germany, August 4-8.

Steyvers, M. & Tenembaum, J. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, Volume 29, 41–78.